Affiliation network: representations and analysis

Pascal Billand, David Frachisse, Nadine Massard

The sixth Framework Program as an affiliation network: Representations and analysis

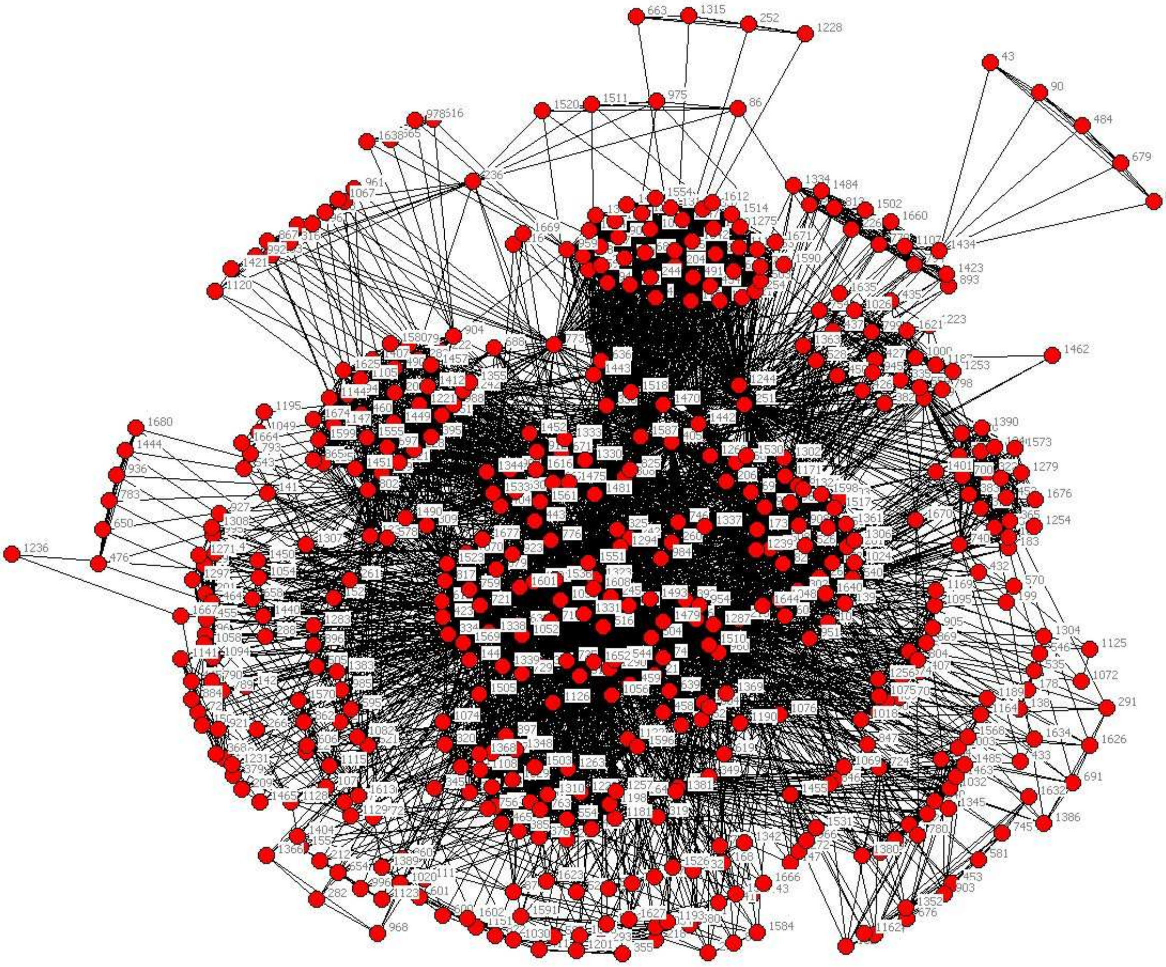
BILLAND Pascal
FRACHISSE David
MASSARD Nadine

CREUSET
Jean Monnet University
Saint-Etienne (France)

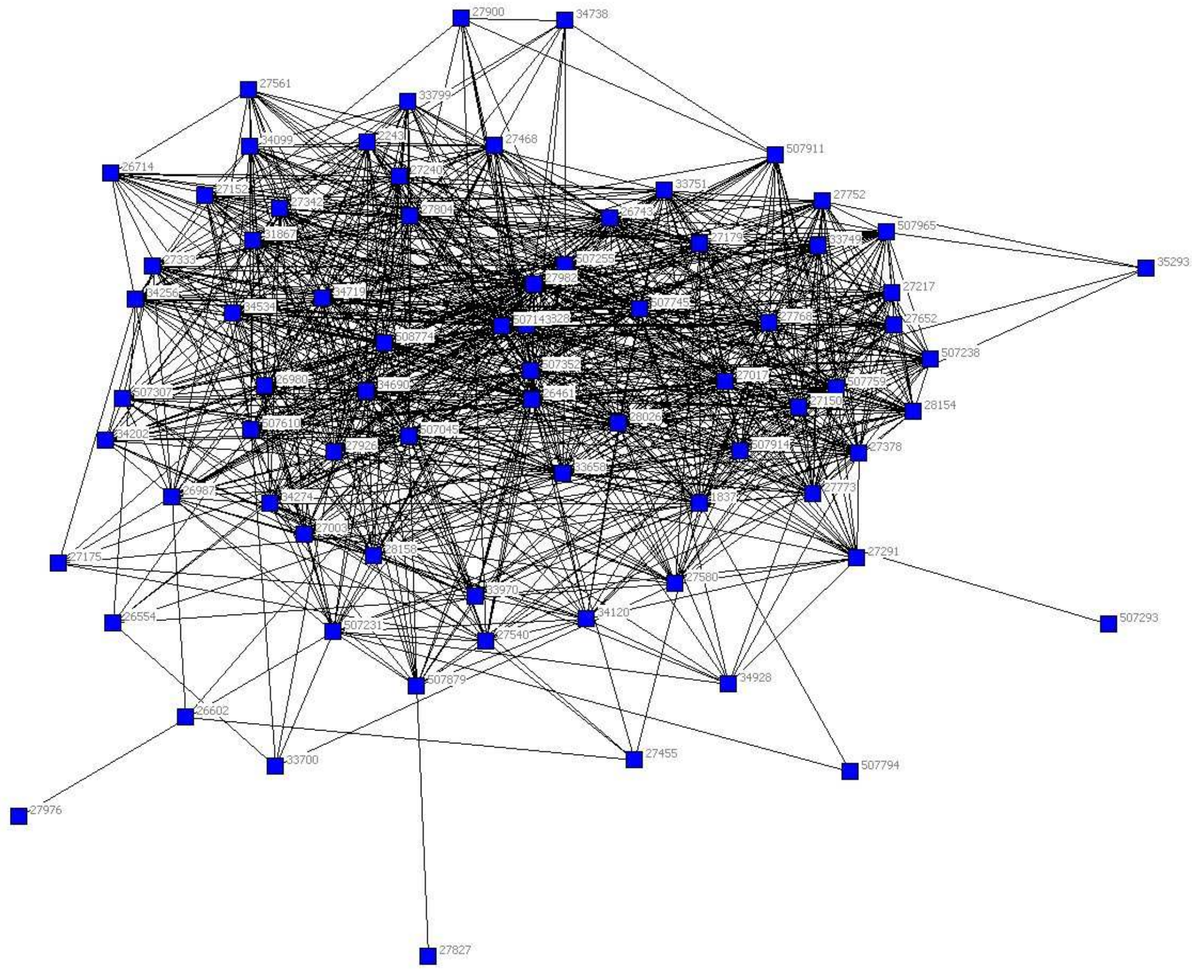
November 2007

Abstract

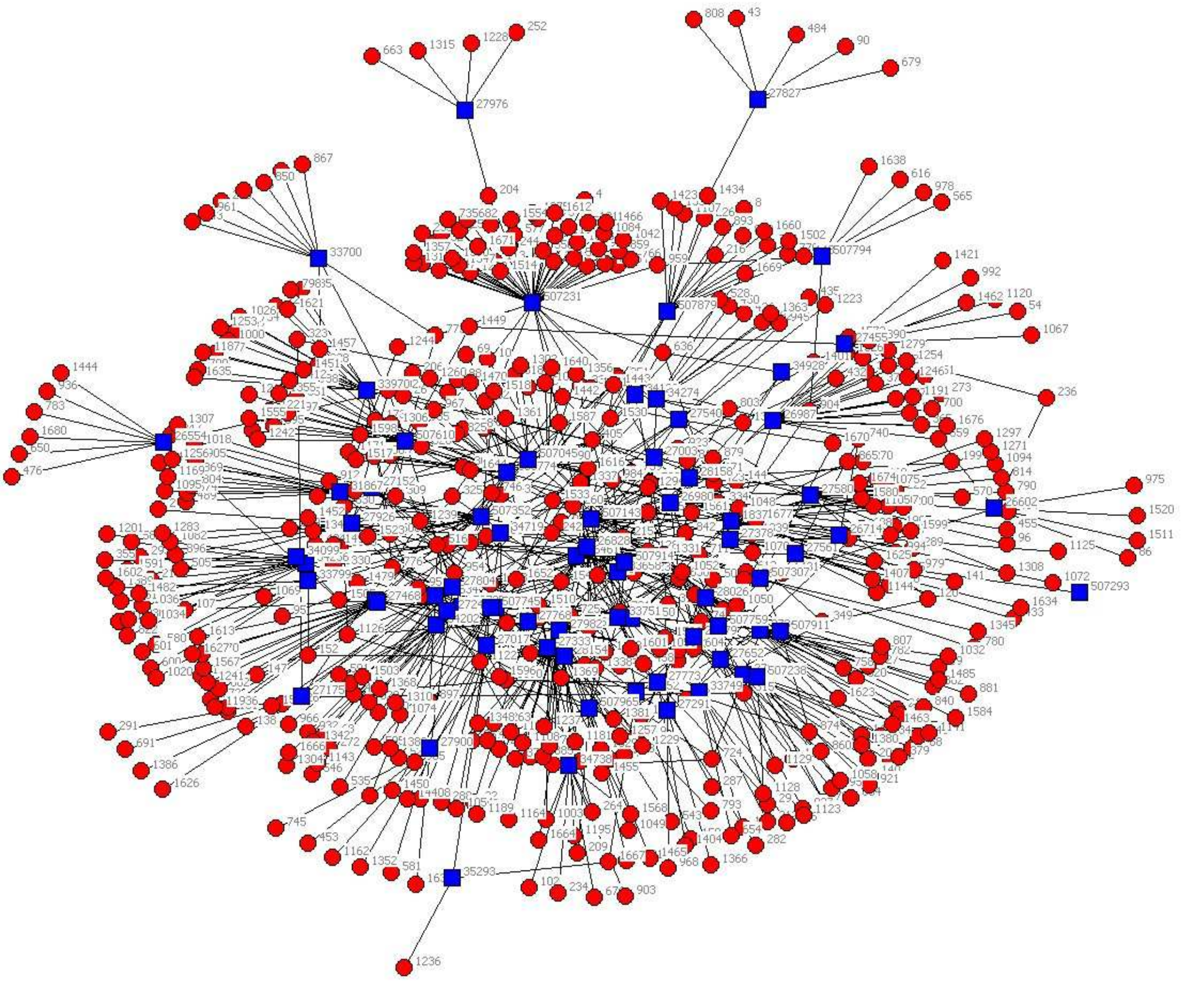
In this paper, we compare two different representations of Framework Programs as affiliation network: "One-mode networks" and "Two-mode networks". The aim of this article is to show that the choice of the representation has an impact on the analysis of the networks and on the results of the analysis. In order to support our proposals, we present two forms of representation and different indicators used in the analysis. We study the network of the 6^{th} Framework Program using the two forms of representation. In particular, we show that the identification of the central nodes is sensitive to the chosen representation. Furthermore, the nodes forming the core of the network vary according to the representation. These differences of results are important as they can influence innovation policies.



Graph 1. Network of agents of the 6th Framework Program (one mode representation). The agents are represented in red. The numbers in the graph correspond to an agent.



Graph 2. Network of events of the 6th Framework Program (one mode representation). The projects are represented in blue. The numbers in the graph correspond to a project.



Graph 3. Two-mode network of the 6th Framework Program. The projects are represented in blue and the agents in red. The numbers in the graph correspond to an agent or to a node.

1 Introduction

A growing literature is emerging on R&D (Research and Development) collaborations between agents (corporations, research centres, universities, ...). These collaborations form research networks. Whilst some papers compare the networks of different industrial fields (Hagedoorn, 2002), others compare the networks with a same industrial field at different points in time (Hagedoorn and Roijakkers, 2006). A part of this literature concerns the networks of collaboration resulting from research projects (Breshi and Cusmano, 2004; Malerba, Vonortas, Breshi and Cassi, 2006; Barber et al., 2006; Roediger-Schluga and Barber, 2006, 2007).

These studies have focused on the R&D networks as they emerge from the EU Framework Programs for Research and Technological Development. One reason for this growing interest is the launching of the European Research Area (ERA) initiative by the EU Commission seven years ago, placing the 6^{th} FP as the main public policy instrument on the way towards ERA. Following the most recent theoretical and empirical debates on collective innovation and R&D networks, the European Commission has clearly emphasized networking in its new objectives. Indeed, Breschi and Cusmano (2004) remark that with the 6^{th} FP as a starting point, policy actions are to be more focussed on identifying crucial nodes and networking centres of excellence, that would represent the backbone of a truly European Research Area. Numerous studies by economists have recently demonstrated the impact of network structures on individual behaviour and economic outcomes. In the case of collaborative R&D networks, for instance, it is worth noting that network structures can influence competition between firms on markets (collaborations causing either a reduction in costs or enhanced quality), the results in terms of technological development (cooperation here is a means of learning, enhancing and orientating the production of new knowledge), or diffusion of innovation.

Improving our understanding of how the networks form and evolve in response to external stimuli is consequently of great importance for designing, implementing and assessing new policy measures that aim to create and deepen the European Research Area. However, applying the network analysis to the context of the European Framework Programs gives rise to numerous theoretical as well as empirical challenges.

As stated by Jackson (2005) the models of network formation arise primarily from two sources: the random graph literature (and the subsequent statistical physics literature) and the economics literature (building on game theory tools). The former responds to the question as to how networks are formed themselves. It starts from the characteristics of the observed networks and highlights the stochastic mechanisms of their formation. The latter responds to the question as to why networks emerge insisting on the individual incentives to create links and enabling conclusions in terms of efficiency of equilibrium networks.

Considering the specific case of R&D collaboration networks formed by the projects proposed to the European Commission within the Framework Programs, we still need to define the relevant theoretical framework. Indeed, the assumption made in the literature based on non-cooperative game theory is that agents (notably firms) form collaborative links by pairs. It is however, not really the case in this context, since agents decide to participate in projects with a set of other agents (firms and other institutional agents). It should be noted that neither can the literature concerning the formation of coalitions be used as firms can participate in different projects with different sets of agents (firms and institutional agents), and can therefore belong to different coalitions (it follows then that the collection of the projects is not a partition of the set of firms).

Moreover, predictions from these different models often take the form of specific network structures such as small-worlds, complete or star networks which are much less complex than what is observable in the real world. Considering the complexity of the European R&D collaboration networks and the associated role of public incentives and individuals maximizing behaviour, it would certainly be worth combining the random and the game theory approaches to networks, as suggested by M. Jackson. Random graph theory could indeed help us to define the impact of public incentives rules upon the global sets of possibilities within the networks while game theory analysis could help us to understand the determining factors for strategic individual decisions within this set of possibilities.

Given our applied perspective however, these theoretical stakes are also closely linked to such empirical challenges as those due to the measurement on data. Within the theoretical framework, standard indicators of network structures and of the positioning of agents are generally used. These are derived from graph theory and social network analysis: indicators of density, distance, centrality, clustering... However, applying these measures to real data often appears difficult and involves methodological choices. So far, few empirical analyses aiming at confronting theoretical construction on the formation of networks with empirical data have been proposed. When they are, R&D collaborations are generally constructed upon data on co-authorship or co-invention. More recently, some studies have focused on networks formed by R&D projects or R&D strategic alliances. Such networks are affiliation networks since they are composed of a set of agents and a set of projects. Authors however, usually choose to transform these networks into "one-mode networks" in order to represent the pattern of relations between actors. Nevertheless, using "one-mode network" in order to represent networks of research projects tends to make one neglect the role of projects per se and hence, may lead to an important loss of information and bias in the analysis, especially when considering the economic dynamics that we are trying to pinpoint through such an analysis of network.

We wish to point out here that there is a lack of understanding of the tools and methodology used for empirical analyses of networks and of their properties. For example, there are numerous methods of measuring the centrality of a node within a network. What is often not recognized is that the formulae for these different measures make implicit assumptions about the manner in which agents interact and flows occur within a network (Borgatti, 2005). These measures clearly identify different facets of a node's role within a network. Underlying hypotheses however, are rarely presented and consequently the interpretation of these measures are difficult to assess.

In this paper, we focus on the methodological implications of choosing a "one-mode network" representation of R&D collaborative networks instead of a "two-mode network" representation. Initially considering the economic questions which underlie the network analysis of European R&D projects, we demonstrate the non-neutrality of the choice of representation. Results concerning the identification of the central players, in particular, differ depending on the chosen representation (one-mode versus two-mode) which may strongly influence the conclusions in terms of public policies.

This paper is organized as follow. Section 2 presents the two approaches of affiliation networks and their representations. In section 3, we show how the chosen representation can impact the analysis of the general properties of the network. In section 4, our interest lies at an inferior level: the identification of the central nodes within the network while, in section 5, we present the core-periphery approach. To conclude, in section 6, we present some limitations of the paper and suggest leads for future works.

2 What kind of network is the sixth European Framework Program of RTD?

2.1 A brief presentation of the European Framework Programs

As collaborative networks, the Framework Programs (FPs) are of particular interest to study. They were created by the European Union (EU) in order to foster the research and development (R&D) cooperation between agents in R&D in different industrial fields and they now constitute its main means of increasing R&D activity. In order to see their project funded, the agents who are involved in a project have to fulfil a number of conditions such as a minimal number of partners, from at least three different countries,... Different funding instruments are proposed such as Network of Excellence (NOE) or Integrated Projects (IP),..., and the agents choose the instrument according to the aim of the project.

If the choice of the representation of the network can influence the conclusion of the analysis, it is likely to influence the direction of the future innovation policy of the EU. It may therefore be interesting to study the difference in the results from network analysis according to representation.

In this paper, we restrict our attention to the 6^{th} FP and more precisely to the field of nano-technology when applied to IST (Information Society Technology) [1].

2.2 The sixth Framework Program as an affiliation network

Affiliation networks are composed of two modes. The first mode is the set of agents N. The second mode is the set of events M. The number of agents in the network is n and the number of events is m. An agent can take part in one or more events. The 6^{th} FP can be seen as an affiliation network. The FPs participants constitute the set of agents and the projects funded constitute the set of events.

We can represent an affiliation network by an affiliation matrix A of size $n \times m$. We denote an element of A by a_{ij}, with $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The coefficient a_{ij} takes the value 1 if the agent i takes part in the event j

[1] We thank the ANRT (Association Nationale de la Recherche Technique) for providing us with the data on the sixth Framework Program.

and 0 otherwise.

$$a_{ij} = \begin{cases} 1 \text{ if } i \text{ participates in } j \\ 0 \text{ otherwise} \end{cases}$$

Each row of A gives all the events in which an agent is taking part. The sum of a row of the affiliation matrix gives the number of events in which an agent is involved. Each column gives all the participants in an event. The sum of a column gives the number of participants in this event.

2.3 "One-mode networks"

Agents and events can be represented separately. More precisely, each mode can be represented by a network and the other is used to set the links of the network. For instance, if we want to represent the network of agents, we use the participations of the agents in the events in order to draw the links between agents. With this approach, we observe only one mode at a time. Graphically, these networks are represented by unipartite graphs.

2.3.1 Networks of agents

A network of agents is formed by a set of agents N and a set of links L_N. Between two agents, we have a direct link if the two agents have at least one event in common.

We build the adjacency matrix of this network X^N. The matrix X^N is a squarred matrix of size $n \times n$. An element of X^N is noted x_{ik}^N with $i,\ k \in N$. The coefficient x_{ik}^N takes the value 1 if i and k have at least one event in common and if $i \neq k$ and takes the value 0 otherwise.

$$x_{ik}^N = \begin{cases} 1 \text{ if } i \text{ and } k \text{ have at least one event in common and if } i \neq k \\ 0 \text{ otherwise} \end{cases}$$

We can build a valued matrix X_V^N. For instance, if two agents have two different events in common, then we can set the value of the coefficient x_{ik}^N to 2 instead of 1. The valued matrix is equal to:

$$X_V^N = A.A'$$

with A being the affiliation matrix and A' the transposed matrix. The diagonal of X_V^N gives the number of events for each agent. If we want to consider only the relations between agents, we have to recode the diagonal and substitute each value in the diagonal with 0. The matrix X^N can be obtained from the valued matrix X_V^N. If the coefficient is 1 or exceeds 1, the coefficient of the matrix becomes 1 in the non-valued matrix. The diagonal of X^N is equal to 0 because we have no reflexive ties.

The network of agents of the 6^{th} FP is given in appendix (Graph 1).

2.3.2 Networks of events

We can use the same method in order to build the network of events. This network is formed by a set of events M and a set of links L_M. Two events have a direct link if at least one agent takes part in these two events.

We build the adjacency matrix of the network entirely composed by events. We note this matrix X^M. An element of X^M is noted x^M_{jl}. A coefficient of this matrix is:

$$x^M_{jl} = \begin{cases} 1 \text{ if } j \text{ and } l \text{ have at least one agent in common and if } j \neq l \\ 0 \text{ otherwise} \end{cases}$$

As for the matrix entirely composed of agents, we can obtain a valued matrix. For instance, if two events have two agents in common, then we can set the value of the coefficient concerning these events to 2 instead of 1. This matrix is noted X^M_V and it is equal to:

$$X^M_V = A'.A$$

with A the affiliation matrix and A' its transposed. The value in the diagonal gives the number of agents per events. If we solely wish to consider solely the relations between events, we have to recode the diagonal as for the adjacency matrix of events. From this matrix, we can obtain the adjacency matrix X^M using the same method as for X^N.

The network of events of the 6^{th} FP is given in appendix (Graph 2). We will show that it is often useful to build the network of events.

2.4 "Two-mode networks"

By definition, affiliation networks are "two-mode networks", the two modes being the set of events and the set of agents.

In order to build the adjacency matrix X^{NM} of the "two-mode network", we can use the matrix affiliation A. This adjacency matrix is built in a different way to the adjacency matrix of the "one-mode network". This matrix is equal to:

$$X^{NM} = \begin{pmatrix} 0_{(n\times n)} & A_{(n\times m)} \\ A'_{(m\times n)} & 0_{(m\times m)} \end{pmatrix}$$

with A being the affiliation matrix, A' its transposed and 0 the zero matrix. In brackets, we can see the size of the matrix which composes X^{NM}.

We use a bipartite graph in order to represent the two modes within the same network. The aim is to observe the events which make the links between agents and conversely. Indeed, in the bipartite graph, there are no direct links between two agents or between two events. Links exist only between agents and events.

Graphically these networks are represented by bipartite graphs. These graphs are built from their adjacency matrix X^{NM}. The network composed of agents and events of the 6^{th} FP is given in appendix (Graph 3).

2.5 Advantages and drawbacks of the two possibilities

The representation of affiliation networks as "one-mode networks" simplifies the analysis in the sense that the network contains only one type of agent. The main advantage of considering only one mode is to improve the readability of the links connecting the elements of a same mode. Indeed, with this representation, there is no intermediary within the relation. For example, we can easily identify agents who have at least one event in common which is useful when studying the relation between the same people at different points in time. At each point in time, we build an affiliation network but the events can be different. The events therefore become less important in the sense that the set of events is different for each period. One example is that of scientific co-authorship. Each year, we can build the network of co-authors, the events being the papers. For each period the events are different although the authors who composed the network may be the same. What matters here is the decision to cooperate with someone whatever the events concerned. Underlying this representation however, there is a central assumption which is that all the authors of a paper work with each other. The analysis may be biased as the links are only potential relations and not real relations. In other words, we can assume that if two agents participate in a same event then there exists a positive probability that these two agents establish relationships with each other. This probability however may still be small. This may especially be the case within a very large R&D project including a great number of agents.

The main interest of representing affiliation networks by "two-mode networks" is to keep the two sets in the graph and to study interactions between these sets.

"Two-mode networks" allow us to retain all the information available in the affiliation matrix. Indeed, in a "one-mode network" of agents, we do not know the event(s) that links two agents together. If three agents are directly connected in the graph, we do not know whether these agents are taking part in the same events or in different events. For instance, suppose that agents 1, 2 and 3 are connected in the "one-mode network". We do not know if these three agents are taking part in the same project or whether agents 1 and 3 are taking part in a project A, agents 1 and 2 in a project B and agents 2 and 3 in a project C. Indeed, these different situations are represented in the same manner in the "one-mode network" whereas the two-mode networks do not have the same affiliation matrix. Studying "two-mode networks" allows us to maintain all the necessary information.

Moreover, in a "two-mode network" links between agents are also considered, although not directly but rather through events. Even though the "two-mode network" expresses a potential it still gives a clear expression of this potential: the projects. Indeed, considering the projects as events limits the bias of the analysis in the sense that the events play a major role in the graph and can be considered as intermediary agents. Furthermore, the projects can be seen as a space within which agents create new knowledge. The basic decision here is that of participating in a project, or the number of projects an agent decides to participate in, thereby determining the structure of the partnerships.

Additionally, this information can be useful for future works concerning the determinants of collaboration between agents especially econometric studies (Autant and al., 2007). In particular, we can study event-by-agent relations or agent-by-agent relations (if we make the strong assumption or if we have more information concerning the way agents interact within the events). We can use variables relating to the agents but also to the events. We can consider the size of the events (as we know, the number of participants in a collaborative research and development project is a determining variable for the European Commission), the participations of an agent in one same event at different points in time, the number of times that two agents participate in the same event. These latter elements give key information concerning the relation between the public policies influencing the dynamics of projects and the individual's decisions influencing the participation of agents and thus the collaboration with others.

Lastly, we can consider these two representations as complementary. Indeed, we can start by studying "two-mode networks" and then move on to a study of “one-mode networks”, particularly agents networks, in order to obtain complementary information. It should be noted that it is always possible to transform “two-mode networks” into “one-mode networks” whereas the opposite is not possible.

The network analysis is based on the study of the general properties of the network (of its graph), the analysis of the place of agents within the network through centrality indicators and the partition of nodes in some sets with particular properties. We will show in the following section that the choice of the representation is important as it changes not only the graph but also the analysis and the conclusion of the study. Moreover, the value of different indicators, particularly centrality indicators, depends on the representation. Finally, the initial choice may condition the final result of a study.

3 The impacts of the choice of the representation on the general properties of the sixth FP network

3.1 Definition of network properties

It is worth noting that the number of components in a network is insensitive to the choice of the representation of the network since it is the same whatever the representation.

By contrast, choosing to represent an affiliation network as a bipartite graph instead of a unipartite graph changes the meaning and the interpretation of some indicators, such as density, distance or diameter. It is important to take into account the partition of the nodes into agents and events on the one hand and the affiliation between the two sets on the other hand.

Concerning density, the partition and the affiliation influence the number of links within the network and the number of potential links, thereby influencing the density of the graph. Indeed, this density is equal to the number of links in the network divided by the number of potential links. In a “two-mode network”, the agents are linked only to events and inversely. Thus, the density of a network agent-by-agent is not comparable with the density of a network agent-by-event even if the both networks come from the same affiliation network. In the “two-mode network”, the density can be considered as an average rate of participation of the agents in the events. In a “one-mode network” of agents, the density refers to the potential of acquaintances realized.

As for the concept of distance, the maximal possible distance between two nodes within the same network depends on the nature of the nodes, on their number and the number of events. This maximal possible distance influences some indicators such as centrality indicators (closeness centrality, betweenness centrality,...). In the following table, we present the differences between the measure with each representation.

Bipartite graph	$m > n$	$m = n$	$m < n$
Maximal distance between two events	$2n$	$2n - 2$ ou $2m - 2$	$2m - 2$
Maximal distance between an event and an agent	$2n - 1$	$2n - 1$ ou $2m - 1$	$2m - 1$
Maximal distance between two agents	$2n - 2$	$2n - 2$ ou $2m - 2$	$2m$

Unipartite graph	$m > n$	$m = n$	$m < n$
Maximal distance between events	n	$n - 1$ ou $m - 1$	$m - 1$
Maximal distance between agents	$n - 1$	$n - 1$ ou $m - 1$	m

We can not compare the average distance between all the pairs of nodes and the diameter in the two representations. Indeed, the indicators of "two-mode networks" are per force higher than the indicators of "one-mode networks". In "two-mode networks", the minimal distance between two agents is two as there is necessary an event which plays the role of intermediary. In "one-mode networks", the minimal distance between two agents is one.

In the next subsection, we apply these concepts to the 6^{th} FP.

3.2 General properties of the network of the sixth FP

Let's us now consider the specific case of the 6^{th} FP. The table below gives some aggregate statistics concerning the network formed by the 6^{th} FP.

Networks	Only agents	Only events	Two modes
Agents	533	-	533
Events	-	75	75
Links	6635	1053	853
Potential links	141778	2775	39975
Density (%)	4.680	37.946	2.134
Components	1	1	1
Diameter	5	4	10
Average distance	2.381	1.696	4.593

Table 1: General properties of the network of the 6^{th} FP.

In this table, there is only one similarity between the two representations: the number of components. As was mentioned in the previous section, the representation does not change the connectivity of the network as the number of components is identical whatever the representation.

By contrast, it may be observed that the number of links as well as the potential links and the density depend on the representation. Authorities should keep this dependence in mind when they wish to set a quantitative target, in terms of a density threshold for instance.

Concerning density, whether we represent the network as a "two-mode network", or as a "one-mode network" and only look at agents, we can conclude that the affiliation network is weakly connected. We reach the opposite conclusion however if we represent the network as a "one-mode network" and look at events instead of agents. The conclusion from these findings differs greatly with respect to the success of EU in creating an ERA. More precisely, we may think that the final objective of the UE is not the FP *per se*, but is in fact the transfer of knowledge between participating agents in the FP. In other words here the FP is a way of inducing agents to meet each other and enabling them to exchange knowledge that they master. In this respect, it appears that this objective is poorly met, as each agent is linked to less than five agents on average. On the contrary, one might think that the EU regards the FP as a centrepiece from the point of view of creation and diffusion of knowledge. Where the participation in a project is crucial for accessing new knowledge, we may observe that each agent benefits from a small part of the total knowledge created within the sixth FP, as each agent only takes part, on average, in around two percent of all the projects. However, we may consider that projects give way to externalities by means of knowledge flows between projects. More precisely, supposing that an agent takes part in two projects, this agent then can be the carrier of knowledge transfers between these two projects. These transfers may benefit agents who participate in these projects. In which case, we must attach importance to the density of events, as it assesses the ability of projects to benefit from the knowledge created by each participant. We observe that the density of events is very high as each project has access to knowledge externalities from approximatively four projects out of ten on average. Here, we can assume that a set of agents, who take part in several events, serves as bridges in the process of disseminating knowledge within the network by creating a number of connections between projects.

We will now consider two distance properties, namely the diameter and the average distance. Looking at the table, we observe that the distance in the "one-mode network" is very low with compared to the "two-mode network": 2.381 versus 4.593 which might influence the conclusion in terms of connectiveness. We can ask what the most appealing indicator is. If we consider that the projects play the role of intermediary agents which are necessary to communicate within the network, we must use the average distance of the"two-mode network". This is the case for instance when agents participating in a project must give their agreement should another partner wish to transmit information to agents outside the project. It is also the case if we consider that projects are places where agents co-construct collective tacit knowledge. Here projects appear as necessary intermediaries in the sense that if one agent wishes to transmit information to an agent from another project he must ask for the consent of or needs the active participation of the other agents involved in his project. By contrast, if we consider that knowledge spillovers are mainly of an informal nature, then we must take into account the average distance in the network of agents. In the relevant literature, projects are often consider as not determining *per se* because the knowledge spillovers are supposed to flow through agents.

Likewise, the choice of the representation has an impact upon the microeconomic properties of the networks. The specific economic role of the central actors or key players on the one hand and the core-periphery structure, on the other hand, are often presented as two common organizing features of many economic networks [2] (Hojman and Szeidl, 2005). These features are behind the numerous connections between economic theory and social network analysis that we have observed recently and which contribute to the use of social network indicators within economic models. They also represent the principal elements of the measure of cohesiveness through networks which can echo the main objectives of the European Commission when implementing FP. We decided therefore to focus our analysis on the measures of centrality that can summarize in particular a node's contribution to the diffusion of knowledge through the network and hence to the cohesiveness of the network, on the one hand, and on the measure of the core-periphery structure on the other hand. This is the purpose of the two following sections.

[2] Especially networks of scientific collaborations (Newman, 2004; Goyal et al., 2006).

4 Individual properties: who are the central nodes?

In the previous section, we observed that the general properties of the network can change according to the representation. We focus now upon individual properties of the network, through centrality indicators, and show that the results of the analysis also depend on the representation.

4.1 Centrality indicators

Numerous definitions of centrality have been proposed under different classifications in the literature on social networks under different classifications (Faust, 1997; Borgatti and Everett, 1997). If we consider that one of the main objectives of the European Commission through the FP is to enhance the diffusion of knowledge through a high level of connectivity between agents and that the role of key actors is to attract and connect new actors within the ERA, then two of these indicators are particularly relevant for our study [3]. Each of these indicators is associated to a specific definition of centrality (Faust, 1997):

- An agent is central if he is active within the network. In order to observe the activity of each agent, we use degree centrality.
- An agent is central if he is linked to central agents. We compute eigenvector centrality to evaluate the centrality.

The first indicator was echoed by Freeman L.C. (1979). The second indicator is a centrality measure proposed by Bonacich P. (1972a, 1972b). Bonacich applied these indicators to "one-mode networks". Faust K. (1997) used these indicators to study "two-mode networks". In addition to the two indicators above, we introduce a new indicator, the strengthened degree indicator, which has some relevance in our study.

In the following section, we define each indicator. We then apply these indicators to the 6^{th} Framework Program in the next section and show the differences between the results concerning the assessment of nodes centrality in the network.

[3] Two other centrality indicators are often used in network analysis: the betweenness centrality and the closeness centrality (Freeman, 1979). We have not used these indicators because these two indicators are not concerned with the activity of the nodes. The former concerns the capacity of a node to avoid the communication between other nodes. The latter concerns whether or not the node is linked in an efficient manner to all the other nodes in the graph.

4.1.1 Degree centrality

Degree centrality is the simplest centrality measure. This indicator measures the centrality of an agent according to his number of links in the graph. The higher this indicator is, the more central is this agent.

One way of interpreting the measure in terms of knowledge diffusion would be to consider that diffusion occurs only through direct links or such processes as specific co-construction of knowledge which imply learning process and cannot diffuse beyond partners directly involved in the construction. Thus, degree centrality measures this capability of nodes to directly influence the integration of agents within the network so that they can benefit from knowledge diffusion and from the research activity of their partners.

In a "one-mode network" representation, we calculate separately the degree of agents and the degree of events.

For agents, we calculate the degree on X^N with the diagonal values equal to 0. The degree of an agent i is the number of different agents who take part in the same events as agent i. It is equal to:

$$C_D^N = X^N.\mathbf{1}$$

with $\mathbf{1}$ a column vector filled entirely with one of size $(n \times 1)$. The degree centrality of an agent i is noted $C_D^N(i)$.

We proceed identically for the network which is composed entirely of events. The degree of an event j is the number of different events which share an agent in common with the event j. In order to calculate this indicator, we use the adjacency matrix X^M with the diagonal values equal to 0. The vector of degree centrality of the events is equal to:

$$C_D^M = X^M.\mathbf{1}$$

with $\mathbf{1}$ a column vector entirely filled with one of size $(m \times 1)$. The degree centrality of an event j is noted $C_D^M(j)$.

In a "two-mode network" representation, the degree of an agent is the number of events in which this agent takes part and the degree of an event is the number of agents who participate in this event. We use the adjacency matrix X^{NM} to compute the degree of each node. The vector gives the degree of the two kinds of nodes and is equal to:

$$C_D^{NM} = X^{NM}.\mathbf{1}$$

with $\mathbf{1}$ a column vector entirely filled of one of size $((n+m)\times 1)$. The degree centrality of a node k is noted $C_D^{NM}(k)$.

4.1.2 Eigenvector centrality

The eigenvector approach aims at measuring the centrality of a node in taking into account the centrality of the other nodes. More precisely, for a node i, its centrality depends on its neighbours centrality and through their neighbours depends upon the nodes with whom it is indirectly linked (Bonacich 1972, 1978, 1991).

This measure is based on the idea of indirect influence. Thus, even if a node influences just one other node, should the latter subsequently influence many other nodes (which themselves influence still more others), then the first node in that chain is highly influential. The eigenvector centrality assumes that knowledge is able to move simultaneously via unrestricted walks rather than being constrained by geodesic paths only. Hence, the eigenvector centrality measure seems well suited for such processes as the diffusion of knowledge externalities within networks.

The centrality of a node is proportional to the centrality of the other nodes and the strength of the link between the nodes. We study a particular case because the strength of the ties is 1 or 0. In the following, we present the concept only for the "two-mode network". Let $C_E(i)$ be the centrality of the node i and x_{ij} the element in the adjacency matrix which reports the relation between i and j. The measure of the centrality of i is:

$$C_E(i) \approx C_E(j) \times x_{ij}^{NM}$$

In order to calculate the centrality of each node within the network we have to solve a system of linear simultaneous equations. This system can be expressed as a problem of eigenvalues and eigenvectors associated to the eigenvalues. The solution is given by:

$$XC_E = \lambda C_E$$

with X the adjacency matrix, λ the high eigenvalue (according to the Perron-Frobenius theorem, this value allows to have an eigenvector with only positive entries) and C_E the vector of centrality which is the eigenvector associated to the eigenvalue λ.

4.1.3 Strengthened degree centrality

We propose another indicator in order to improve the analysis. We have seen that the degree of a node is a simple measure which indicates the number of links of the node. However, this measure does not take into account the centrality of the nodes with whom a node is linked. Next, we have presented the eigenvector approach, which considers that the centrality of a node depends upon the centrality of their neighbours (partners). However, this measure can be difficult to calculate when the number of nodes becomes large and it is generally restricted to the main component.

Our indicator attempts to bypass these difficulties encountered by the eigenvector approach. More precisely, the aim of our indicator is to calculate the centrality of a node taking into account the degree of this node and the degree of the other agents within the network. We require an indicator which is simpler than the previous one and which allows us to calculate a measure of centrality even when the network is not entirely connected and whatever the number of nodes in the graph. Our indicator is based on the degree centrality and the geodesic distance (the shortest path) between the nodes. This indicator name is "Strengthened degree centrality". The centrality of an agent is noted $C_{SD}(i)$.

We can give the following interpretation to this indicator. Referring to the law of gravitation, the attraction of a distant body is supposed to be equal to its mass weighted by a decreasing function of its distance. Therefore considering the activities to be reached as being attractors and the distance as the impedance function (friction of distance), several accessibility indicators have been elaborated upon by geographers (Schuermann and al., 1997; Vickerman and al., 1999). They have been used by economists through potential function to measure market potential or accessibility to knowledge in a geographical context with a simple inverse function to model the distance decay effect. The idea here is to consider that the centrality of an agent is not only linked to his own activity (degree) but also depends upon his capacity to access the activity of all other agents within the network. The simplest measure of distance through relational networks is the geodesic distance between two agents and this distance is considered as infinite when there is no path between two agents. To calculate this indicator, we use the degree centrality vector C_D and the geodesic distance matrix G. We denote by g_{ij} an element of G which is the geodesic distance between the nodes i and j. The centrality of a node i depends on the centrality of all the other nodes within the network but the influence is inversely proportional to the geodesic distance between i and all the other nodes within the network. With this indicator, a node is central if it is connected to other central nodes or if it is located near other central nodes, even though it is indirectly connected to the latter.

In order to calculate the strengthened degree indicator, we make the following assumption: the distance between two nodes, not directly or indirectly linked, equals infinity. If we want to apply our indicator to all the networks we have to transform the matrix of geodesic distance. Let us denote by G^* the transformed matrix and by g^*_{ij} an element of this matrix. The diagonal of G is entirely composed of zeros. In order to take into account the degree of a node we add one to all the elements of the matrix G. In this way, we do not change the relative distance between nodes. Next, we take the inverse of each element obtained in this manner. An element of G^* is noted g^*_{ij}. So also byusing also the inverse function as the decay [4], the vector of centrality C_{SD} is equal to:

$$C_{SD} = G^*.C_D \text{ with } g^*_{ij} = \frac{1}{1 + g_{ij}}$$

The centrality of a node i is:

$$C_{SD}(i) = \sum_{j \in N} g^*_{ij} \times C_D(j)$$

So there is no decay for the personal activity of the agent considered.

We note that if a node i is not connected with a node j, the centrality of i does not depend on the degree of j. This allows us, indirectly, to take into account the size of the component as the centrality of an agent depends only on the degree of the nodes in the same component. In order to be central, a node has to be linked with other active nodes and to be in the main component, everything else being equal (in order to be connected with many other nodes). This indicator can be applied identically to the two representations. In the case of "two-mode networks", the centrality of an agent is higher if he takes part in large events which are composed of agents who have a high degree and who also take part in large events,...

Additionally, such an indicator can lead way to different assumptions notably in terms of influence. We have already mentioned that different ways of modelling the decay are possible. We can also introduce thresholds thereby restricting the distance of influence between two nodes, assuming that the centrality of a node depends only upon nodes localized at a distance inferior

[4] For a discussion on the relevance of such a form see Kwan (1998).

to this threshold. If we study knowledge spillovers for example, one usually assumes that a given node can not benefit from the knowledge of another node beyond a given distance. Furthermore, our indicator allows us to discriminate more precisely between the agents. Indeed, one drawback of the degree centrality is the fact that many agents have the same degree indicators whereas they do not have a symmetrical position within the network. Our indicator allows us to rank the nodes more precisely in the network in terms of centrality.

Throughout this subsection, we have presented centrality indicators which report the position of the nodes in the network. A better understanding of these indicators and of their properties is essential to the network analysis from an economic perspective. Central agents are often considered to be best performers in economic networks (Hojman and Szeidel, 2005) and they exercise an attractive power over the other participants in the networks. Agents may strategically wish to work with the more central agents, as these latter can be more powerful, more skilled,... Thus identifying the key players is a central question for economists especially when analysing collaboration networks. This can help us to understand individual collaborative behaviours and assess the role of public incentives policies on cooperation. Henceforth, it may be useful to assess the extend to which the results concerning agents centrality depend on the representation of the network and on the indicators.

4.2 Who are the central nodes in the sixth Framework Program?

In this section we show that the choice of the representation of the network and the centrality indicators can affect the identification of the central nodes in the 6^{th} FP.

In order to compare the two representations, we calculate centrality indicators for all the nodes. Note that the use of a cardinal approach to compare the value of the indicators in the two representations is irrelevant. Henceforth, we use an ordinal approach and rank the nodes. More precisely, in the case of the "two-mode network", we propose two rankings: one for the agents and one for the events. We next compare the rank of a set of nodes within the two representations. We calculate the three indicators introduced above in order to see whether or not the rank of a node for an indicator changes with the representation. We obtain a rank for each node and each indicator. We then compute the average rank for each node. We denote this latter

$RANK$. For each ranking, we present only the ten most central nodes (see Tables 2, 3, 4 and 5 in Appendix).

At this point we should make a few general remarks. Firstly, we can see that the ten most central nodes (agents or projects) are not the same according to the representation. Between the two agents rankings, we have only two agents in common and we have only five projects in common between the two projects rankings. This shows that the choice of the representation influences the ranking of the more central agents. We will explain this difference later.

Secondly, the rank of each node changes not only according to the representation but also according to the choice of the indicators.

Thirdly, in the "two-mode network", the ranks of each node are relatively more stable whatever the centrality indicators than in the "one-mode network". This may suggest that the "two-mode network" is less sensitive to the choice of the indicators. However, we must be cautious with such a proposal as it is based on one example only.

Finally, when we take the absolute value of the difference between the rank of an agent in one representation and the rank of this node in the other representation and we add these values for each node, we observe that the total for the strengthened degree is weaker than the total for the eigenvector centralities and the degree. This may suggest that our indicator is less sensitive to the representation than the eigenvector approach.

We now explain more precisely the difference between the results and give some interpretations for each indicator according to the representation.

Degree centrality is the number of links that a node has within the network, but this number does not have the same meaning according to the representation. Recall that, in the "one-mode network", we set links between all the agents who have at least one project in common. Henceforth, the degree of an agent i is the number of different agents who take part in at least one event in common with i. However, in the case of the FP, we can suppose that there exists a threshold for the number of agents in a given project beyond which no real collaborations exist between all the partners. Indeed, in the case of the 6^{th} FP, 32 projects out of 75 have more than ten agents. It is difficult to assert that effective ties exist between all the agents in such projects, even though they can communicate for data transfer or a minimum of consultation or coordination. Concerning the degree of a project j in the "one mode network", it measures the number of different projects which have at least one agent in common with j. The centrality indicators of the projects are rarely studied. Generally, studies on affiliation network focus only on the network of agents. Yet, it is important to observe the network of events in order to find the central projects as those which have the highest potential to diffuse knowledge on the one hand and benefit from knowledge spillovers on the other hand.

Should we wish to look at the agents who connect projects and hence permit the flow of knowledge, then we must represent the two modes within the same network. The degree calculated on the "two-mode network" offers immediately interpretable information. In the "two-mode network", the degree of an agent is the number of projects in which he takes part and the degree of a project is the number of agents in the project. Hence, degree centrality gives two different and complementary set of information for each representation. If we want to calculate a more refined measure we have to combine the representation in "two-mode network" with eigenvector centrality or strengthened degree. This combination enables us to take into account the direct links between agents and projects as well as indirect links between two nodes of a same mode. For instance, for a given agent, we can take into account the number of projects and the number of potential partners. This way, we consider the two modes in the centrality of a node.

Eigenvector centrality, or strengthened degree centrality, are based on the same intuition: the centrality of a node depends on the centrality of the other nodes. The major difference of the "two-mode network" when compared to the "one-mode network" is that the centrality of a node depends on the two kinds of nodes: the agents and the projects. In this case, the projects are considered as intermediary agents which allow connections between agents. In the case of the "two-mode network", taking part in many projects or to having many participants is insufficient to be central. In order to be central, an agent should not only take part in many projects but the projects should also contain central agents, *id est* agents who take part in central projects. In the "one-mode network", in order to be central, an agent must be connected to other central agents. The size of some projects which by definition create a lot of links may therefore induce bias.

In order to show that the choice of the representation is not neutral, we take as an example the project number 507231. This project contains 53 agents. Among them 42 take part in this project only. By definition, these 42 agents have a degree centrality of 52. This explains why in the representation of the affiliation network as a "one-mode network", these 42 agents are ranked in the top even though they take part in only one project. We can argue that the centrality property of an agent must take into account not only the number of potential partners, but also the number of projects in which he takes part. It reflects the skill of the agent to work in different fields, to gather knowledge coming from different research topics so avoid becoming dependent on only one project related to his research activity,... In order to take into account the number of projects per agent and the potential partners, we must use eigenvector or strengthened degree indicators within the "two-mode" representation, and following these indicators, the agents who only take part in the project 507231 get down at the bottom of the ranking.

Now, if we consider the rank of the project itself we remark that the project 507231 is more top-ranked in the "two-mode network" than in the "one-mode network" for the degree centrality. This can be explained as follows. This project is the biggest project within the network and the degree centrality in "two-mode network" is equal to the number of agents that it contains. As it is the biggest project, it therefore occupies first place in the ranking. We could ask why it is not also first in the "one-mode network"? The answer is that many agents who take part in this project participate only in this project. they therefore do not allow for the creation of connections with other projects. In this sense, we can not say that this project is really central. As for the agents, we recommend the use of another indicator rather than the degree centrality and use of the "two-mode network" so that the centrality of a project is a function of the centrality of its agents.

The European Union might wish to identify central nodes allowing an efficient transmission of knowledge either because they greatly participate in the FP or because they occupy a strategic position (being a cutpoint for instance[5]). The European Union might wish to identify the central agents of the network of the FP in order to anticipate who will be the preferred partners for the next FP. Furthermore, the EU might want to find central projects in order to know which kind of instruments are central in the network for connecting agents, which fields in the nanotechnologies are central,... Depending upon the representation, the answer to these questions will be different.

4.3 What types of node are central nodes?

4.3.1 The agents

In the Framework Program, we consider five kinds of agents: the agents from the Higher Education (HE), the research centres (RES), the small and medium firms ($IND-SME$), the big firms (IND) and another type of

[5]A cutpoint is a node which is deleted from the network, increases the number of components in the network

agents (OTH) in which are gathered all the other types of agents. The table below gives the number of agents of each type.

Kinds of agents	Number of agents
Higher Education (HE)	136
Research Centre (RES)	100
Small and Medium Firms ($IND - SME$)	134
Firms (IND)	121
Others (OTH)	42

The question now is to determine which types of agents are the most central in the network of the 6^{th} FP. The first remark is that even if the ten most central agents differ according to the representation, none of them are from small or medium firms or belong to the kind OTH. The kind OTH is composed of associations, federations... who take part in one project only. They are not really research agents. The fact that no small and medium firms are top ranked might be due to their poor means. More than 90% of them participate in one project only.

Secondly, we observe that the top ten differ according to the representation. In the case of a "two-mode network", the top ten central agents are 5 firms, 3 research centres and 2 agents from Higher Education. By contrast, in the "one-mode network", 6 agents from Higher Education, 3 research centres and only 1 firm compose the top ten. According to the representation, the conclusions concerning the kinds of institutions represented by the most central agents change and hence reveal behavioural differences. The importance given to the very big projects within the one-mode representation contributes to designating the partners who most frequently participate in these big projects as central agents hence favouring partners from Higher Education. On the contrary, the importance given to the role of the intermediaries between projects within the two-mode representation tends to favour the designation of firms as central agents.

More generally, we can compute an average rank by type of agent in order to know which are globally more central. For each type of agents, we can compute an average measure for each indicator. Next, for each indicator, we rank the types of agents from one to five. Finally, we obtain a general ranking for each type of agent. See tables 6 and 7 in the appendix.

In both cases, the most central agent are on average the agents from Higher Education and the research centres. We can assume that this is due to the fact that the FPs are a real opportunity for these agents to obtain fundings and to develop collaboration with other types of agent especially firms. By contrast, the small and medium firms are the least central category although the FPs also offer them a real opportunity to collaborate. We can assume that some research centres and agents from Higher Education have greater means and are therefore able to participate in more projects than the small and medium firms, the CNRS in France as one example. Indeed, we have seen that more than 90% of the small and medium firms participate in only one project against 70% for the research centers and 60% for the agents from Higher Education. Firms are not central in the network of the FP. This could be due to the fact that the firms, notably the more innovative ones, develop industrial partnerships outside the FP and so essentially focus their cooperation within the FP looking for research centres or agents from Higher Education partners, in order to benefit from other kinds of knowledge. Additionally, small and medium firms can be competitors of other firms on the market. Industrial partners might therefore prefer to choose collaboration with public or research centres in order to avoid transmitting knowledge to competitors. Finally, it is not only the chosen representation that influences the centrality ranking of each type of agents, the choice of the indicators is also influential.

4.3.2 The projects

In the 6^{th} Framework Program, the agents have the choice between five kinds of project: Coordination Action (CA), Integrated Projects (IP), Network Of Excellence (NOE), Specific Support Action (SSA) and Specific Targeted Research Project (STP).

Kinds of projects	Number of projects
Coordination Action (CA)	3
Integrated Projects (IP)	25
Network Of Excellence (NOE)	4
Specific Targeted Research Project (STP)	37
Specific Support Action (SSA)	6

As for the agents, we analyse the kinds of the ten most central projects. In the “two-mode network”, we find 8 Integrated Projects and 2 Network of Excellence. By contrast, in the “one-mode network”, we find 6 Integrated projects, 2 Network of Excellence and 2 Specific Targeted Research Project. The presence of IP and NOE is normal as these two types of project are the main instruments of the EU for fostering of competitiveness of the EU in area of R&D. By definition, they are big projects in terms of number of agents. Furthermore, the most central are composed of central agents thus obtaining the highest centrality indicators. As we can see with project 507231, however this is not always the case. The presence of two Specific Targeted Research Projects out of the ten within the one-mode representation tend to proves that the centrality of IP within the two-mode networks is not only due to their size but also due to their capacity to connect large projects to each other through intermediary agents.

With regard to the analysis of the centrality of each type of project, we use the same ranking method as for the agents. Results are in Tables 8 and 9 in the appendix.

The first remark is that the general rank is the same whatever the representations. The IP and NOE projects are on average the most central projects. This is not surprising considering the particular interest and the targets given to these instruments by the EU. The size effect suffices to differentiate these instruments in terms of centrality whatever the representation. The second remark is that, as for the agents, the choice of the indicators changes the rank of the instruments. Once more this is a demonstration of the methodological consequences of the choice of indicators. There are two possibilities for improving the robustness of the results. On the one hand, we can use several indicators and compare the results or we can insist on particular properties and choose the most appropriate centrality indicator regarding these properties, on the other hand.

5 Core-periphery approach

We now introduce the core-periphery approach and apply it to the 6^{th} FP.

5.1 Definition

The core-periphery approach is a method which partitions nodes into two sets: central nodes and peripheral nodes (Borgatti and Everett, 1997, 1999, 2000). This method uses a numerical method to find a partition of nodes which is as close as possible to an idealized matrix. This idealized matrix is divided into four parts. The upper left-hand corner is the core and ideally the density between nodes is 100%. The density in the lower right-hand corner is equal to 0%. We do not consider the value of the density in the two other corners. There exit various assumptions concerning the upper right-hand and the lower left-hand corners, we present only one of them.

100%	-
-	0%

The numerical method consists of performing a high number of iterations. For each iteration, a fit is calculated. The fit function is the correlation between the permuted data matrix (obtained at each iteration) and an ideal structure matrix consisting of 1's in the core block interactions (upper-left corner) and o's in the peripheral block interactions (lower-right corner). In this case, we use the correlation between the adjacency matrix and the idealized matrix. The higher the fit is, the best the partition is right. One measure for comparing the real matrix to the ideal is as follows:

$$\varphi = \sum_{i,j} x_{ij}\gamma_{ij}$$

$$\gamma_{ij} = \begin{cases} 1 \text{ if nodes } i \text{ and } j \text{ belong both to the core} \\ 0 \text{ if nodes } i \text{ and } j \text{ belong both to the periphery} \\ . \text{ otherwise} \end{cases}$$

with φ the fit, x_{ij} the element of the adjacency matrix used and γ_{ij} the element of the ideal matrix. With this measure, we treat the off-diagonal values as missing data.

In "two-mode networks", the core is a partition of agents CO_N connected to many events belonging to a partition of events CO_M and, simultaneously, this partition of events CO_M is composed of events which are connected to many agents who belong to the partition of agents CO_N. The rows of the matrix are agents and the columns are events. In "one-mode networks", the core is a partition of agents or events who are close to each other.

When using this method, we must be careful when interpretating of the results. We should recall that the iterations to find the partition of nodes cease when the procedure reaches a maximum. This maximum can be local, however. Moreover, the core-periphery structure of the network can be determined by the adjacency matrix with which we start the procedure. In other words, for different starting points, we can obtain different results. thus, we have to launch the procedure with different adjacency matrixes as starting points and keep the result that is most frequently obtained.

5.2 The same affiliation network but different cores

The core-periphery approach allows us to strengthen the definition of centrality in the sense that to be central a node must have a high centrality indicator and also be localized in the core of the network too. In this subsection, we apply the analysis to the 6^{th} FP and show that the partition core-periphery partition depends upon the representation of the network used.

The core of an affiliation network, especially in the case of the FP must be composed of agents and of projects. As a way towards the ERA, the FPs are supposed to help improve the cohesiveness of the R&D networks and the integration of marginal actors within the processes of creation and diffusion of knowledge. Developing core-periphery measures is a way of identifying the agents who are highly connected (those who are in the core) and of assessing their role in establishing relations with partners from the periphery. Within European R&D networks however, we must not forget the projects which act as intermediaries. Indeed, the European Commission does not validate bilateral links between agents but a set of agents grouped in projects which have particular properties. It is the affiliation of the agents to this project which makes the agents central. The application of the core-periphery approach to "two-mode networks" slightly modifies the definition of the core. The set of agents in the core is defined by the set of projects in the core and conversely.

There is a high number of nodes in the FP, so it is difficult to present the total core-periphery partition. We present tables which sum-up the partition. We compare the number of nodes in the core and in the periphery in the two representations for each set of nodes. We note CO the core and P the periphery. The indice indicates the set of nodes studied (N for agents and M for events) and the exponent the number of modes studied at the same time. For example, the agents localized in the core of the "two-mode network" is noted CO_N^2. We obtained the following results concerning the 6^{th} FP network.

Agents	CO_N^2	P_N^2	Total
CO_N^1	254	223	477
P_N^1	27	29	56
Total	281	252	533

Projects	CO_M^2	P_M^2	Total
CO_M^1	15	24	39
P_M^1	13	23	36
Total	28	47	75

The agents and the projects in the core differ according to the representation. This result requires cautious however, as it depends on the capacity of the algorithm to find the core. We have seen that the algorithm can cease at a local maximum. Thus, the core obtained by the algorithm is not necessarily the set of nodes that maximize the density in the upper left corner and minimize the density in the lower right corner. If we repeat the process however, with different starting points, this problem may be avoided.

As done for the centrality indicators, we observe the difference in the kinds of agents and the kinds of projects which form the core between the two representations. In the core of the "one-mode network", the five types of projects have at least one project in the core while in the core of "the two-mode network" only three types of project (NOE, IP and STP) are represented. Moreover, the projects NOE and IP are more highly represented in the core of the "two-mode network". They represent more than 75% of the core in the "two-mode network" against only 50% in the other case.

We might ask what kind of measure is relevant. Should we want to prove that these types of projects play a central role within the network of the 6^{th} FP, then the statistics of the "two-mode network" are more relevant. For instance, take a project p which contains many agents, all of them but one however, the agent i, take part only in p. Agent i participates in many projects. Through the agent i, the project p is connected to all the projects in which the agent i takes part. The density between p and the other projects of i is one. Thus, the set formed by p and all the projects of i meet the properties of the core while p is not really central in the sense that only one agent make the link between p and the rest of the network. This simple example suggest that it may be relevant to use the "two-mode network" in order to identify the core.

It should also be noted that the number of nodes in the core is always lower in the case of the "two-mode network". The core of the "two-mode network" contains 28 projects and 281 agents. By contrast, the core of the network of events contains 39 projects and the core of the network of agents contains 477 agents. Hence, the application of the core-periphery approach allows us to restrict the core and to be more selective in the identification of the major nodes of the network.

On the whole, keeping the information on the projects in the network and applying the two-mode representation allows us to avoid the bias which could be created by the biggest project. This is especially relevant in the case of FP. Indeed, when we create a link between each agent in the same project we create a set of agents who respond perfectly to the definition of the core. As all the agents in a given project are connected in "one-mode networks", the density between these agents is one. The algorithm can identify this subset of agents as representing the core of a “one-mode network”. One might argue that the problem is the same for “two-mode networks”. If we consider a project, all the agents in the project are linked to this project, so the density is one. The difference lies in the fact that, in the case of the “two-mode networks”, we can easily see that the core suggested by the algorithm corresponds only to one project.

The density matrixes are:

- For the “two-mode network”:

0.039	0.019
0.027	0.010

- For the “one-mode network” of agents:

0.105	0.012
0.012	0.023

- For the “one-mode network” of events:

0.714	0.307
0.307	0.148

Concerning the density matrix, we should be careful when interpreting the results. Indeed the matrix shows weak results in the case of the 6^{th} FP. The only network that presents good results is the network of projects. In this case, we have a high density in the upper right-hand corner and a low density in the lower left-hand corner. Another limitation to the application of the core periphery method in the “two-mode network” is the fact that two nodes which are symmetrical in the network (in the sense that these two nodes are substitutable without changing the properties of the network) can be localized differently: one in the core and the other in the periphery. For instance, in project 507231 which we studied before, some agents (those who take part in only one project) are in the core whereas the others are in the periphery. Yet, all these agents are symmetrical within the graph in the sense that if we permute two of them, the properties of the graph do not change and the individual properties of the two nodes are identical.

In few words, this application, as reported here, suggests that the choice of the representation changes the results of the analysis. We have said in the previous section that the EU might want to keep the most central agents for future Framework Programs and thereby favouring projects with these agents in the 7^{th} FP. The EU may favour a set of agents, however who presents many connections in order to accelerate the process of knowledge creation and knowledge transfer. Not being able to directly solicit bilateral relations, it is essential for the European Commission to measure the impact of the projects themselves on the definition of the core. This contributes to reinforcing our argument in favour of the two-mode representation.

6 Conclusion

In this paper, we have considered two methods of representation of the 6^{th} FP. Most of the studies only look at the one-mode representation and focus their attention on the links between agents. The application to the R&D European network as they emerged from the FPs allows us to show that the "one-mode network" representation is likely to distort the analysis. The two modes are necessary to correctly analyse the network as an agent influences the centrality of an event and conversely. In this paper, we have shown that the study of the "two-mode network" and the study of the network of events bring new and complementary information to the study of the network of agents. We show that the conclusions that emerge from the analysis regarding the efficiency of the network are different depending upon the representation. Indeed, we have seen that the general properties of the network (density, average distance, diameter) are different from one representation to the other. The conclusions in terms of efficiency of the network can change between the representations. Furthermore, in the case of FPs, we can see also that the central nodes differ according to the representation. Finding central nodes can be highly relevant for the EU in order to know which agents in the network are important for the efficiency of the network. If we regard only the results on the "one-mode network", the conclusion may be biased in the sense that some agents who are considered as central agents in reality only take part in one project. The identification of central nodes differs also according to the choice of the indicators. In order to find central nodes in term of activity and influence on the rest of the network, we recommend the use of the eigenvector centrality or the strengthened degree in the "two-mode network". This combination enable us to take into account the projects in which an agent takes part as well as the potential partners of this agent. Moreover, we have also shown the influence of the chosen representation when using the core-periphery approach. Once more we recommend to use the core-periphery approach in the "two-mode network" as the projects are necessary for an agent in order to be central and conversely. Finally, with “two-mode networks”, it is always possible to study “one-mode networks” while the opposite is not true. Using “one-mode networks” however is a solution, when we are sure that the agents in the events have a real contact such as in co-authorships with few co-authors for example.

In order to correctly describe the properties of the network and to find central agents, we should cross analyses using the different networks. Besides, we also have to cross the divers centrality indicators and the core-periphery approach. For instance, we can say that an agent is central should he takes part in a minimum of projects, have a minimum number of potential partners, have high centrality indicators whatever the representation and is in the core of the two representations.

An insight into future work would be to build a valued matrix instead of a binary matrix as an adjacency matrix. This would allow us to consider the repetition of the tie between agents in the case of the “one-mode network”. If we take into account the repetition of the link between a given pair of nodes in the adjacency matrix, the conclusion of the analysis may also change. It will be interesting to study the difference in the results that we obtain if we analyse the same network when in one case we consider a binary matrix and in the other case a valued matrix. Further empirical analyses are also envisaged using econometrics technics in order to better understand the determinants of the relations between agents and between agents and events. We can focus our attention on the number of times the agents are in the same events. We can attempt to explain this by their past affiliations, for instance in the 5^{th} FP, variables relative to the agents and variables relative to the events.

References

[1] Allen M.P. (1974), "The Structure of Interorganizational Elite Cooptation: Interlocking Corporate Directorates", *Amercian Sociology Review*, **39**: 393-406.

[2] Autant-Bernard and al. (2007), "Social Distance versus Spatial Distance in R&D Cooperation: Empirical Evidence from European Collaboartions Choices in Micro and Nanotechnologies", *Papers in Regional Science*, **86**, 3, August: 495-519.

[3] Ballester C. and al. (2006),"Who's who in networks. Wanted: the key player", *Econometrica*, **74 (5)**: 1403-1417. 176-185.

[4] Barber M.A., Krueger T. and Roediger-Schluga T. (2006), "The network of EU-funded collaborative R&D projects", *Physical Review E*, **73**, 036132 (arXivv:physics/0509119).

[5] Bonacich P. (1972a),"A Technique for Analysing Overlapping Memberships", *Sociological Methodology*, edited by Herbert Costner, 176-185.

[6] Bonacich P. (1972b), "Factoring and Weighting Approaches to Status Scores and Clique Identification", *Journal of Mathematical Sociology*, **2**: 113-120.

[7] Bonacich P. (1978), "Using Boolean Algebra to Analyse Overlapping Memberships", *Sociological Methodology*, **9**: 101-115.

[8] Bonacich P. (1991), "Simultaneous group and individual centralities", *Social Networks*, **13**: 155-168.

[9] Borgatti S.P. (2005), "Centrality and network flow", *Social Networks*, **27**: 55-71.

[10] Borgatti S.P. and Everett M.G. (1997), "Network of analysis of 2-mode data", *Social Networks*, **19**: 243-269.

[11] Borgatti S.P. and Everett M.G. (1999), "Models of core/periphery structures", *Social Networks*, **21**: 375-395.

[12] Borgatti S.P. and Everett M.G. (2000), "Peripheries of cohesive subsets", *Social Networks*, **21**: 397-407.

[13] Borgatti S.P., Everett M.G. and Freeman L.C. (2002), Ucinet for Windows: Software for SocialNetwork Analysis, Harvard, mA: Analytic Technologies.

[14] Breschi S. and Cusmano L. (2004), "Unveiling the structure of an European Research Aera: Emergence of oligarchic networks under E.U. Framework Programmes", *International Journal of Technology Management*, **27**: 747-782.

[15] Breiger R.L. (1974), "The duality of persons and groups", *Social Forces*, **53**: 181-190.

[16] Faust K. (1997), "Centrality in affiliation networks", *Social Networks*, **19**: 157-191.

[17] Freeman L.C. (1979), "Centrality in Social Networkd: Conceptual Clarifiaction", *Social Networks*, **1**: 215-239.

[18] Goyal S., Van der Leij M. and Moraga J.L. (2006), "Economics: An emerging small world?", *Journal of Political Economy*, **2**: 403-438.

[19] Hagedoorn J. (2002), "Inter-firm partnerships: an overview of major trends and patterns since 1960", *Research policy*, **31**: 477-492.

[20] Hagedoorn J. and Roijakkers W. (2006), "Inter-firms R&D partnering in pharmaceutical biotechnology since 1975: Trend, patterns, and networks", *Research Policy*, **35**: 431-446.

[21] Hojman D. and Szeidl A. (2005), "Core and Periphery in Endogenous Networks", KSG Faculty Research Working Paper Series, RWP06-022, Harvard University.

[22] Jackson M. (2005), "The economics of social networks, Proceedings of the 9^{th} World Congress of the Econometric Society", edited by Richard Blundell, Whitney Newey, and Torsten Persson, Cambridge University Press.

[23] Jackson M. and Wolinsky A. (1996), "A strategic model of economic and social networks", *Journal of Economic Theory*, **71**: 44-74.

[24] Kwan M.-P. (1998), "Space-Time and Integral Measures of Individual Accessibility. A Comparative Analysis Using a Point-Based Framework", *Geographical Analysis*, **30(3)**, 191-216.

[25] Malerba F., Vonortas N., Breschi S. and Cassi L. (2006), "Evaluation of progress towards a European Research Area for Information Society Technology", Report for the European Commission, CESPRI.

[26] Newman M.E.J. (2004), "Coauthorship networks and patterns of scientific collaboration", *National Academy of Science USA*, **101**: 5200-5205.

Appendix

In the following four tables, we substitute the name of the agents and the name of the projects in order to respect data confidentiality. The exponent on the indicators indicates the adjacency matrix used and so the representation used.

Agents	Kind	C_D^{NM}	Rank	C_E^{NM}	Rank	C_{DS}^{NM}	Rank	$RANK$
721	RES	27	1	0,287	1	543,27	1	1
516	RES	24	2	0,282	2	536,83	2	2
215	RES	19	3	0,237	3	500,32	3	3
1052	IND	13	4	0,157	4	474,31	5	4,33
1287	IND	13	4	0,142	6	467,00	6	5,33
405	HE	12	6	0,134	7	479,67	4	5,67
638	IND	10	7	0,151	5	453,83	7	6,33
1510	HE	8	9	0,109	8	451,17	8	8,33
251	IND	9	8	0,083	10	449,34	9	9,00
1294	IND	6	12	0,088	9	436,88	10	10,33

Table 2: Ranking of the ten most central agents in the "two-mode network".

Agents	Kind	C_D^N	Rank	C_E^N	Rank	C_{DS}^N	Rank	$RANK$
251	IND	155	4	0,159	1	5569,67	4	3
1470	HE	116	6	0,155	2	5432,08	5	4,33
1587	HE	89	12	0,149	3	5369,17	6	7,00
1442	HE	90	11	0,148	4	5305,92	9	8,00
336	HE	73	16	0,139	5	5196,75	11	10,67
773	HE	84	13	0,139	5	5120,75	15	11,00
1518	HE	73	16	0,138	7	5142,08	14	12,33
1443	HE	66	22	0,132	8	5040,25	18	16,00
516	RES	251	1	0,092	54	5869,25	1	18,67
636	RES	61	26	0,131	9	4984,83	23	19,33
1244	RES	65	23	0,131	9	4943,25	26	19,33

Table 3: Ranking of the ten most central agents in the "one-mode network".

Projects	Kind	C_D^{NM}	Rank	C_E^{NM}	Rank	C_{DS}^{NM}	Rank	$RANK$
26828	IP	34	2	0,302	1	528,87	1	1,33
27982	IP	27	4	0,18	4	487,94	6	4,67
26461	IP	20	11	0,203	2	514,51	2	5,00
508774	IP	29	3	0,169	5	486,82	7	5,00
507143	NOE	22	9	0,197	3	501,20	3	5,00
507352	IP	22	9	0,144	6	492,85	5	6,67
507255	IP	23	7	0,139	7	486,68	8	7,33
34690	NOE	16	13	0,136	8	497,33	4	8,33
1837	IP	27	4	0,102	10	457,71	12	8,67
507045	IP	15	15	0,121	9	470,73	10	11,33

Table 4: Ranking of the ten most central projects in the “two-mode network”.

Projects	Kind	C_D^{M}	Rank	C_E^{M}	Rank	C_{DS}^{M}	Rank	$RANK$
26828	IP	60	2	0,201	2	945,17	6	3,33
507255	NOE	53	5	0,183	6	888,50	17	9,33
507352	IP	54	4	0,188	5	864,00	24	11,00
27017	NOE	38	18	0,141	18	926,00	8	14,67
34719	IP	41	14	0,148	14	899,33	16	14,67
507745	IP	46	9	0,171	7	847,33	30	15,33
507045	IP	44	11	0,154	13	873,50	23	15,67
26461	IP	63	1	0,205	1	822,17	46	16,00
28026	STP	45	10	0,164	9	850,92	29	16,00
34690	STP	47	7	0,162	10	841,08	35	17,33

Table 5: Ranking of the ten most central projects in the “one-mode network”.

We denote R_D the rank for the degree indicator, R_E the rank for the eigenvector indicator, R_SD the rank for the strengthened degree and R_G the general rank. The exponent indicates the adjacency matrix used and so the representation used.

Kinds of agents	R_D^{NM}	R_E^{NM}	R_SD^{NM}	R_G^{NM}
HE	2	2	1	1
IND	3	3	2	3
IND-SME	5	5	5	5
OTH	4	4	4	4
RES	1	1	3	1

Table 6: Ranking of the kinds of agents in the "two-mode network".

Kinds of agents	R_D^N	R_E^N	R_{SD}^N	R_G^N
HE	2	3	1	2
IND	4	4	4	4
IND-SME	5	5	5	5
OTH	3	1	3	3
RES	1	2	2	1

Table 7: Ranking of the kinds of agents in the "one-mode network".

We denote R_D the rank for the degree indicator, R_E the rank for the eigenvector indicator, R_SD the rank for the strengthened degree and R_G the general rank. The exponent indicates the adjacency matrix used and so the representation used.

Kinds of agents	R_D^{NM}	R_E^{NM}	R_{SD}^{NM}	R_G^{NM}
CA	3	3	3	3
IP	2	2	2	2
NOE	1	1	1	1
SSA	5	4	4	4
STP	4	5	5	5

Table 8: Ranking of the kinds of projects in the “two-mode network”.

Kinds of agents	R_D^{M}	R_E^{M}	R_{SD}^{M}	R_G^{M}
CA	4	3	4	3
IP	2	2	2	2
NOE	1	1	1	1
SSA	3	4	5	4
STP	5	5	3	5

Table 9: Ranking of the kinds of projects in the “one-mode network”.